194-813-SI4-004

EOSDIS Core System Project

DADS Prototype One STK Wolfcreek 9360 Automated Cartridge System Hardware Characterization Report

May 1994

Hughes Applied Information Systems
Landover, Maryland

DADS Prototype One
STK Wolfcreek 9360 Automated Cartridge System
Hardware Characterization Report

May 1994

Prepared Under Contract NAS5-60000

APPROVED BY

Stephen Fox /s/	5/18/94
Steve Fox, SDPS Office Manager EOSDIS Core System Project	Date

Hughes Applied Information Systems
Landover, Maryland

This page intentionally left blank.

Abstract

The Storage Technology Corporation's (STK) WolfCreek 9360 Automated Cartridge System (ACS) was used in the evaluation. It is a recently released new tape library. As compared to STK's earlier 4400 ACS libraries, WolfCreek is a smaller-capacity higher-performance library.

The hardware characterization data obtained consisted of the robotics action time measurement, the data throughput figures, and the cumulative drive mount (or dismount) and communication delay time. These performance figures will be used in calibration of the Data Archival and Distribution system (DADS) component performance modeling.

This evaluation report summarizes the hardware characterization results yielded during prototype evaluation. The chief purpose of the prototyping effort was to evaluate the operation of two Commercial Off The Shelf (COTS) File Storage Management Systems (FSMS): E-Systems Modular Automated Storage Systems (EMASS) Group's FileServ and Convex Computer Corporation's (Convex) UniTree, with a commercial Automated Tape Library (ATL). Hardware results were peripheral to the main goal of the prototype.

Specific emphasis was placed on the investigation of the influence of the FSMS on the data transfer rates. The data transfer rates observed during the use of the Fileserv differed significantly from those obtained during the use of the UniTree. The rates are detailed in the report. The WolfCreek silo under test proved dependable, and was easy to operate.

This page intentionally left blank.

Preface

This document summarizes the results of the hardware portion of the DADS Prototype One performed between September 1993 and December 1993 in the Data Archival and Distribution System (DADS) section of the Science Data Products System (SDPS) within the EOSDIS Core System project.

This is an informal document that is approved at the ECS Office Manager level and does not require Government approval. After this prototype has been completed, a final report will be documented in the Prototyping and Studies Final Report, DID 331/DV3.

For additional technical information pertaining to the prototype, contact Steve Fox, SDPS Office Manager at 301- 925-0346 or via electronic mail sfox@eos.hitc.com.

This page intentionally left blank.

Change Information Page

List of Effective Pages	
Page Number	**Issue**
Title	Original
iii through viii	Original
1-1 and 1-2	Original
2-1 and 2-2	Original
3-1 and 3-2	Original
4-1 through 4-8	Original
5-1 and 5-2	Original
AB-1	Original

Document History			
Document Number	**Status/Issue**	**Publication Date**	**CCR Number**
194-813-SI4-004	Original	May 1994	

This page intentionally left blank.

Contents

Figures

Tables

Abbreviations and Acronyms

1. Introduction

The primary goal of the prototyping effort was to evaluate the operation of two Commercial Off The Shelf (COTS) File Storage Management Systems (FSMS): E-Systems Modular Automated Storage Systems (EMASS) Group's FileServ and Convex Computer Corporation's (Convex) UniTree, with a commercial Automated Tape Library (ATL). The evaluation of the Storage Technology Corporation's STK WolfCreek silo's capabilities was a secondary goal.

The purpose of the hardware evaluation was to briefly examine and time the performance of the main ATL components: robotics and tape drives, and to spend the bulk of the hardware effort examining the dependencies between the use of a specific FSMS and the data throughput rates resulting. Operation in degraded mode was examined as time allowed. No stress testing of the equipment was done due to the time constraints.

The hardware characterization data obtained consisted of the robotics action time measurement, the data throughput figures, and the cumulative drive mount (or dismount) and communication delay time. The data transfer rates observed during the use of the Fileserv differed significantly from those obtained during the use of the UniTree. The rates are detailed in the report. The ease of operation and dependability of the STK WolfCreek silo created a very positive impression with the evaluation team. The published operating benchmarks for the WolfCreek silo were largely confirmed. A strong dependency was noted between the file size and the resulting tape read and write performance. The results of this evaluation augment the work performed in the "STK WolfCreek Beta Summary" study by Gary G. Hull, Hughes STX at Goddard Space Flight Center (GSFC).

Prior to the prototyping effort, STK provided four hours of technical training on the operation of the WolfCreek ACS. Technical support was readily available. The virtually flawless operation of the ATL necessitated very little problem resolution contact with STK personnel during the prototyping period.

The rest of the document contains the hardware evaluation report and a summary of the hardware related issues that arose during the evaluation.

This page intentionally left blank.

2. Hardware Configuration

The WolfCreek ACS configuration used in the prototype normally holds 1000 cartridges. The WolfCreek unit installed for the prototype had only 800 cartridges actually residing in the silo. The reduced capacity was due to a transparent viewport taking up some of the silo's wall space.

The terms ***drive*** and ***tape transport*** must be clarified. In STK terminology a ***drive*** specifies an entire enclosure housing one or more tape transports with the associated interfaces and power supplies. The term ***tape transport*** is, as defined in the glossary of the STK's *Hardware Operator's Guide* document [STK Document number 3086255-01], "An electromechanical device capable of threading tape from a cartridge, moving the tape across a read/write head, and writing data onto or reading data from the tape." In the FileServ logs, as well as in a colloquial usage of the term, ***drive*** refers to the same hardware mechanism that STK specifies as a ***tape transport***. Therefore, subsequently in this paper both ***tape transport*** and ***drive*** terms are used interchangeably to refer to the tape transport.

Figure 2-1 illustrates the hardware configuration. The control path connection (Figure 2-1, path 5-6-3) was via Ethernet.

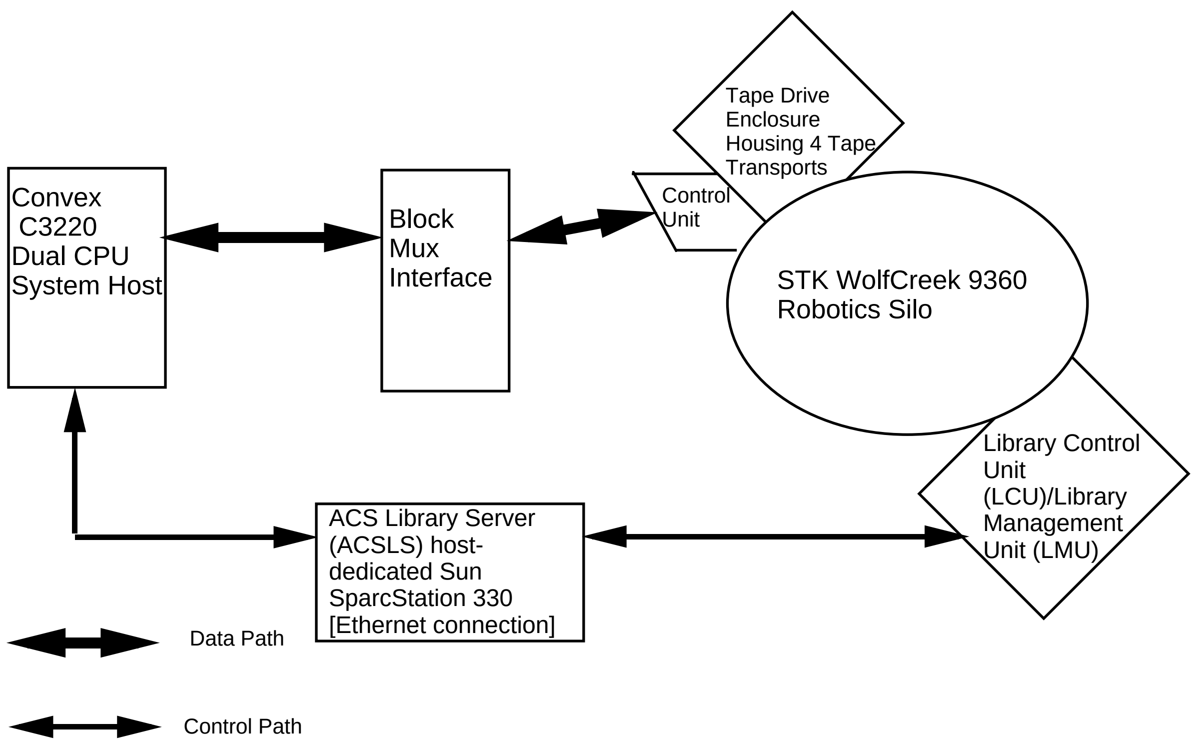

Figure 2-1. Prototype 1 Hardware Configuration

Convex C3220 Dual Central Processing Unit (CPU) System host was configured as follows:

DISK SPACE:	20 Giga Bytes (GB) in 8 disks
OS VERSION:	10.1.2 - initially; 10.1.4 - after 11/10/93
MAIN MEMORY:	2 GB

Storage system:

ROBOTICS UNIT:	STK WolfCreek 9360 Silo
CARTRIDGE FORMAT:	3480
CARTRIDGE CAPACITY:	800 Cartridges
VOLUME CAPACITY:	200 MB/Cartridge
TAPE TRANSPORTS:	4
DATA TRANSFER INTERFACE:	Block Mux

ACS Library Software (ACSLS) HOST: Sun SparcStation 330

ACS System Administrator (ACSSA) SOFTWARE: Release: 4.00 Variant: Sun 2.0

3. Software Component

In the first portion of the prototyping effort the hardware performance was evaluated using FileServ 2.1.6 from the EMASS Group. A software tool was written by Data Archival and Distribution System (DADS) personnel to create the data files for storage and retrieval. Tape drive read and write rates, combined tape transport and communication path delays, as well as the "stop watch" measurement of robotics action time were measured using FileServ.

After the planned tests using FileServ were completed, FileServ was deinstalled and Convex UniTree 1.75.14 was installed in its place. Tape drive read and write rates were the only hardware parameters collected using UniTree.

This page intentionally left blank.

4. Summary of the Hardware Tests

4.1 Test Data

The file distribution listed in Table 4-1 was stored and retrieved to obtain hardware timings. The time intervals were derived by correlating the reported initiation and completion times in the logs. The logs were generated by the FSMS, Convex OS, and the software tool developed by the DADS personnel. Five to twenty measurements in each reported category (depending on log entry availability) were used to compute the average throughput rates for each category of files. The number of measurements used is reported in Tables 4-2 and 4-3.

The granularity of the FileServ FSMS timing logs was 1 second due to the use of the standard Unix timing calls, which resulted in loss of precision in the measurements obtained using FileServ. For example, for a 100 MB file with an average reported time expended on the data read of 37 seconds, the error due to the measurement precision is +/- 0.07 MB/sec (2.6% of the reported value). For a 1 KB file with an average reported time expended on the data read of 1 sec, the error due to the measurement precision is +/- 0.0005 MB/sec (50% of the reported value). A better resolution of the UniTree debug level 4 statements in the *tapemovr.t* logs allowed greater precision on the performance figures reported by Unitree.

Additional data on robotics action was obtained using a stop watch. The video camera mounted on the robot was very helpful in facilitating the stop watch measurements. The remaining portion of this section outlines selected aspects of hardware operation.

Table 4-1. Files stored and retrieved

Class Name	# of Files	File Size(s)	# of Tapes Used
A	40	100 MB	20
B	400	10 MB	20
C	4000	1 MB	25
D	6000	100 KB	25
E	99	100 MB, 1 MB	9
F	908	100 MB, 10 MB	11
G	9210	100 MB, 100 KB	14
H	1040	10 MB	13
I	30	1 KB, 200 MB, 300 MB, 500 MB, 1 GB	34
J	5000	1 KB	8
K	5000	1 KB	4
L	5000	1 KB	5
M	5000	1 KB	4
N	5000	1 KB	4
O	5000	1 KB	5
P	5000	1 KB	4
Q	5000	1 KB	5
R	5000	1 KB	5
T	4500	1 KB	5
Misc.	132	Other Test Files	15
TOTAL	71359	71680 Possible Inodes*	235/600**

Notes: * - Overall number of inodes available in a partition is equal to the total number of files allowed in that partition.

** - 600 is the total number of tapes dedicated for the use by FileServ of which 235 were used.

4.2 Robotics Action and Tape Mount/Dismount

The robot operated consistently and dependably throughout the prototyping period. A physical inventory of the silo contents was automatically performed by the robot every time the silo door latch was re-latched and on power up, on the assumption that a tape position in the silo may have altered.

The average time expended by a combination of a robotics retrieve and drive mount action was 17.14 seconds. This measurement was taken using the time differential from the FileServ FSMS "ETAC" log entry reporting the drive allocation and the "drive in use" report from the FileServ "hist" file. The "stop watch" measurement of robotics action time was 6 seconds on average for a storage or a retrieval operation. The rest of the time, 11.14 seconds, is attributable to the operations internal to the tape transport, such as tape threading and de-threading, and the communication delay in the control path. The communication delay is affected by the control path being on the common Earth Observing System Data and Information System Core System (ECS) network.

4.3 Tape Drive Operation

4.3.1 Data Transfer Rates

The blocking factor used by FileServ in storing to and retrieving from tape was 64 KB. UniTree used a 16 KB blocking factor (15 KB of actual data plus 1 KB for the Unitree headers). According to the information provided by STK technical support personnel, 64 KB is the largest blocking factor, that the STK 3480 tape transports under prototype evaluation can support. Packetization overhead is associated with each block of data. Therefore, both the throughput performance of the drive and the storage capacity of the tape are affected by the blocking factor selected by the FSMS. Table 4-2 summarizes the capacity dependencies of the 3480 tape cartridge (200 MB nominal native capacity).

Table 4-2. Storage Capacity Dependency of the 3480 Tape on the Data Blocking Factor (200 MB Nominal Capacity)

Blocking factor	128 Bytes	1,024 Bytes	8192 Bytes	16 KBytes	32 KBytes	65.5 KBytes
Storage Capacity in MB	9	57.03	173.5	203.15	222	233

The measurements summarized in Table 4-3, Read Throughput Rates, and Table 4-4, Write Throughput Rates, were obtained from the logs generated by each FSMS. An attempt was made to report "pure" read and write transfer rate values that do not include time delays for the robot action, tape mount and dismount, and file seek time.

Using FileServ, performance figures were obtained by averaging out the time intervals computed from the entries in the FileServ logs for each of the size categories. Standard Unix time calls are used by FileServ for the time stamps in the logs. The time expended for a read or a write operation was calculated by subtracting the time reported in the "component [transport] in use" (or "COPY STARTED") entry from the time reported in the "command successful" (or "COPY COMPLETED") entry. The information on a file's relative position on the tape, afforded by FileServ, allowed to isolate the actual read and write times from the file seek delay with some degree of certainty. Using FileServ logs, time measurements taken for file at the start of tape were correlated with the time measurement for the files 2nd from the start of tape, 3rd from the start of tape etc. in order to isolate the file seek time from the actual data transfer rate.

For UniTree the data was produced by averaging out the reported performance figures from the UniTree debug level 4 statements in the *tapemovr.t* logs. It must be noted, that the initial data throughput rates reported by Unitree were surprisingly low. When Convex technical support personnel was consulted about the problem, it was discovered, that the read and write rate calculation algorithm contained an error. That was when the debug checkpoints were turned on to produce the figures shown in this report. No independent verification of the Unitree calculations was made by the prototype team.

The size of the file transferred to and from the tape had a significant effect on the throughput performance. As can be seen in the Tables 4-2 and 4-3, performance for both the read and the write operations drops off sharply for files of 1 MB and below. The reduction in performance can be attributed to the overhead associated with handling of smaller files (such as the file seek time and the time spent on reading, writing or processing of header information associated with each file).

Figures 4-1 and 4-2 show the plots of the data in Tables 4-2 and 4-3 respectively. During the use of FileServ there was a slight performance improvement on the read function for 64 KB - sized files as compared to both 100 KB and 1 KB files. The hypothesis is, that the 64 KB blocking factor used during the read makes the retrieval of a file that is precisely 64 KB in size more efficient. The blocking factor did not seem to favourably affect the writing performance for the 64 KB files due to the overhead associated with the write operation.

The control path was on the common ECS network. The data path was dedicated via the BlockMux interface. Therefore the only influence of the network on the throughput performance data was in reporting delays affecting the "started" and "stopped" entries in the FileServ logs. As compared to the relatively significant measurement error due to the 1 second precision, network delays for the two log entries may be considered negligible. The measurement error also masks the standard deviation for the throughput rate.

Table 4-3. Read Throughput Rates

File Size	1K	64K	100K	1M	10M	100M
FileServe Read Rate in MB/sec	0.001	0.064	0.033	0.5	2.5	2.7027
UniTree Read Rate in MB/sec	0.00118	0.0864	0.133	0.763	1.411	1.651
Number of Samples	5	5	10	10	20	20

Table 4-4. Write Throughput Rates

File Size	1K	64K	100K	1M	10M	100M
FileServe Write Rate in MB/sec	0.0005	0.016	0.025	0.033	1.25	2.5445
UniTree Write Rate in MB/sec	0.00099	0.0583	0.0842	0.477	0.905	1.022
Number of Samples	5	5	10	10	20	20

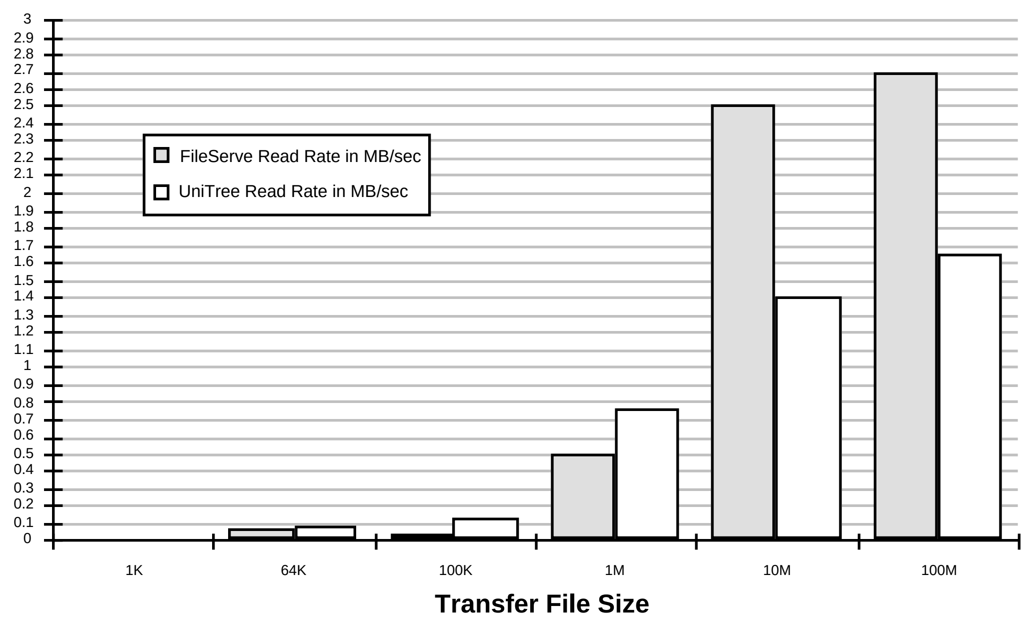

Figure 4-1. Read Throughput Rates

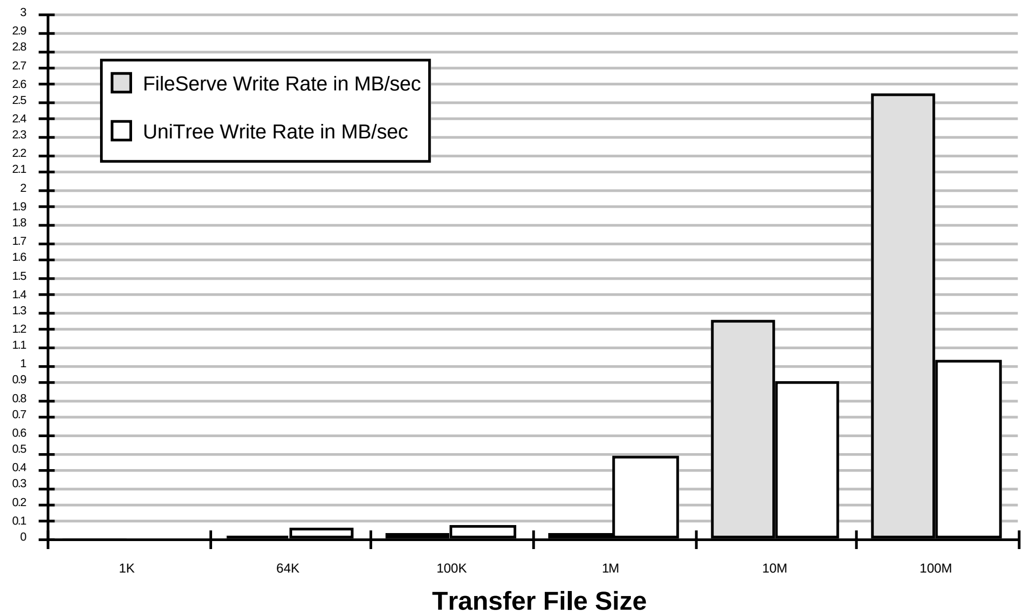

Figure 4-2. Write Throughput Rates

The marginally better performance of the read and write operations for the file sizes of 1 MB and smaller using UniTree is not fully understood. One possible contribution to the performance lag could be the reporting precision of the FileServ logs that has more impact on the timing results for smaller files read over a shorter time interval than it does on the timing of the larger files.

4.4 Problems Experienced During Prototyping

Only one problem during the prototyping period was due to a hardware malfunction. The problem occurred twice during the prototype period. Two different drives missed a grab of a tape leader. The problem is described in more detail in Section 4.3.1 below. STK technical support personnel promptly resolved the first occurrence of this problem and provided instruction, that enabled the prototyping team to resolve the second occurrence of the problem.

The second problem, described in Section 4.3.2, *End of Tape Writing Problems*, while initially attributed to the hardware, proved to be the Convex OS problem and was resolved by an application of the version 10.1.4 of Convex OS.

4.4.1 Missed Tape Leader Problem

During the early stages of prototyping an STK transport "missed" its grab of the tape leader. This resulted in a "tape mounted" report from the drive, but the transport could not be used and the tape could not be dismounted by a *dismount* command with a *force* option issued at the ACS

Library Software (ACSLS) host. STK technical support personnel powered off the transport and manually removed the tape. FileServ was not being tested at the time.

The problem occurred a second time on a different transport early during the FileServ evaluation. At that time it was corrected by DADS personnel. In both cases, the remainder of the silo's tape transports as well as the robotic arm continued to function normally.

STK technical support personnel attributed the problem to the height of the drive arm being slightly out of adjustment. Technical support specialist stated, that the problem of that nature occurred very rarely and, that once the arm is adjusted, the problem should not recur on that drive.

The problem arose in the initial stages of the prototype tests, shortly after equipment installation, and has not occurred again. Therefore, it is anticipated, it would occur rarely in normal operations environment. A problem like that would be flagged as part of the drive availability monitoring done routinely in an operations environment. It would have a minimal impact on a facility operation as a whole (see Section 4.5, Degraded Operations Mode).

4.4.2 End of Tape Writing Problems

Repeated problems writing to the end of the tape were experienced intermittently from 11/2/93 through 11/10/93 during FileServ write operation testing. A drive failure report and a log entry of an "*error positioning to end of tape, ioctl error*" flagged the problem. After five such errors FileServ changed the drive status to "*not available*". In several cases the media was marked suspect by FileServ.

After several consultations with Convex, EMASS, and STK technical support personnel the problems were attributed to the Convex Operating System (OS) version 10.1.2 time-out problem and not to the STK hardware. Application of version 10.1.4 of Convex OS resolved the problem. No further problems writing to the end of tape were experienced.

4.5 Degraded Operation Mode

Problems experienced as described in Sections 4.3.2 and 4.3.3 afforded a natural opportunity to test the hardware in a degraded mode of operation, i.e. with one or more tape drives varied (marked) off-line. No performance degradation was noted in the remainder of the ATL.

Specific degraded mode testing occurred by powering down drives that were not in use. This was done with and without FileServ's knowledge. In each case FileServ continued operations normally. Hardware error reporting was correct. Error reporting was used by both the hardware remaining on-line and the FSMS to allowed for uninterrupted operation of the healthy components.

Tests using one FSMS were considered sufficiently indicative of hardware behaviour. For the results of degraded mode of operation testing using Unitree see the "Prototype One. FSMS Product Operational Evaluation" Paper.

4.6. Ease of operation

Overall, the ease of operation for the STK library was assessed as very good. Error and event logging for the ACSLS is clear and detailed, recording equipment status, such as "door open", as well as errors.

The technical assistance afforded by the STK Technical support personnel was prompt and constructive. Consistent and trouble-free operation of the hardware necessitated very little STK assistance for problem resolution.

5. Conclusion

The results obtained during this prototyping effort were compared with those of an earlier "STK WolfCreek Beta Summary" study performed by Gary G. Hull, Hughes STX at GSFC. It was noted, that the robotics action measurements of 6 seconds were identical in both studies. The read and write data transfer rates noted in the "Beta Summary" diverged from the rates obtained in this study: transfer rates on both the read and the write operation were significantly higher using Cray Data Migration Facility (DMF) for the transfers of 128 KB and less as compared to both Convex UniTree and FileServ. No measurements were reported by the "Beta Summary" for transfers of blocks larger than 128 KB. Therefore, it is not possible to correlate the two studies.

The WolfCreek ACS library was dependable, easy to operate and requiring minimal maintenance. Error and event logging in ACSLS is clear and detailed.

A strong dependency was identified between the file size and the resulting tape input/output performance. Specifically, the throughput rate dropped off sharply for files of 1 MB in size and smaller. A better throughput performance of the drives under UniTree control as compared to FileServ was noted for files of the size of 1 MB and smaller. For file sizes beginning with 10 MB and larger FileServe performance was considerably better, than that of Unitree.

This page intentionally left blank.

Abbreviations and Acronyms

ACS	Automated Cartridge System
ACSLS	ACS Library Software
ACSSA	ACS System Administrator
ATL	Automatic Tape Library
COTS	Commercial Off The Shelf
CPU	Central Processing Unit
DADS	Earth Observing System (EOS) Data and Information System (EOSDIS) Core System (ECS) Data Archive and Distribution System
DMF	Data Migration Facility (Cray)
ECS	Earth Observing System Data and Information System Core System
EMASS	E-Systems Modular Automated Storage Systems
FSMS	File Storage Management Systems
GB	Giga Bytes
GSFC	Goddard Space Flight Center
KB	Kilo Bytes
MB	Mega Bytes
OS	Operating System
STK	Storage Technology Corporation